Presence in Immersive Virtual Environments

Mel Slater and Martin Usoh,
Department of Computer Science, and
London Parallel Applications Centre (LPAC),
QMW University of London,
Mile End Road,
London E1 4NS, UK.

1. Introduction

Immersive Virtual Environments (IVE) provide a tightly coupled human-computer interface: input to the sensory organs of the human participant are directly generated through computer displays, in the visual, auditory, tactile and haptic modalities. The participant operates in an extended virtual space, created by the interaction between the human perceptual system, and the computer generated displays. In this paper we outline some of the results of a pilot experimental study of presence in IVEs. This is a contribution to a project involved in constructing a system for architectural walkthrough, where architects and their clients are able to navigate through and effect changes to a virtual building interior. The results of a long term research effort on interior walkthrough, concentrating mainly on visual realism is presented in (Brooks, 1992). In our work we have concentrated on the interface provided by the virtual environment generator (VEG) to the human user, and initially on the problem of the establishment of the presence of the human inside the VE. Our project partners are concentrating on photo-realistic image generation for building interiors (Harriss and McKellar, 1992).

Related work is reported in (Slater and Usoh, 1992, 1993a, 1993b). The current paper includes additional material and some conclusions to date. In Section 2 we outline our current understanding of the factors contributing to presence. In Section 3 we describe our pilot experiments. We divide the factors contributing to presence into exogenous and endogenous. The exogenous factors are those provided directly by the VEG itself and we discuss the results relating to these in Section 4. The endogenous factors are subjective, helping to explain people's differing responses to the same experiences. We discuss these subjective endogenous factors in Section 5 followed by conclusions in Section 6.

2. Presence

In this research we are concerned with the extent to which a human subject allows him/herself to be convinced while experiencing the effects of a VEG that s/he is somewhere other than where s/he physically is - that "somewhere" being determined by the images, sounds and physical sensations provided by the VEG to the person's senses. This is similar to Ellis' definition of "virtualisation" (Ellis, 1991). Note that this involves a suspension of disbelief; the person truly "knows" that s/he is in one place while experiencing inputs to his/her sense organs that to some extent provide the experience of being somewhere else.

How can we know when someone is "present" in a VE? This question has no commonly accepted answer. However, we would expect such an answer to include the following (Held and Durlach, 1992; Sheridan, 1992; Loomis 1992a,b):

(a) The participant can verbally report the extent of such presence;
(b) An external observer can note the reactions of the participant while experiencing events or performing tasks in the IVE, and note how closely their responses are to those observed in everyday reality;
(c) An external observer can note the extent to which the participant fails to respond to events in everyday reality while immersed in the VE;

and we would add:
(d) The person can later report the sense of having been somewhere other than where they really were at the time.

0-7803-1363-1/93 $3.00 © 1993 IEEE

Points (a)-(c) could apply equally well to someone very absorbed in a movie - they might identify strongly (a form of presence) with the situation of one or more of the actors; experience strong emotions and visibly react to events within the movie (for example, experience vertigo, pain, fright, horror, love); and be so absorbed that they become totally unaware of the noises and sights of their real surroundings. Point (d) is less likely to apply in this case. In response to a question "Where are you?" - even the most absorbed movie watcher is likely to answer "In the cinema", whereas the participant in an IVE is likely to describe the scene being presented to his sense organs - in particular his eyes - by the VEG.

We distinguish exogenous and endogenous factors responsible for determining the extent of presence. By exogenous factors we mean those supplied directly by the VEG itself, i.e., the signals provided to the sense organs of the human participant. These signals are a function of the hardware and software of the VEG, including how the system responds to the actions of the human by altering subsequent signals, thus forming a feedback loop between the human subject and the VE system. Such exogenous factors, to be detailed below, provide the *necessary* conditions for a sense of presence in the VE.

The recent literature on presence has identified the following exogenous factors likely to be important in determining the degree of presence (Held and Durlach, 1992; Loomis, 1992a; Sheridan, 1992; Heeter, 1992):

2.1.1 high resolution information being presented to the appropriate sensory organs with the information received through channels to all sensory organs describing a consistent world;
2.1.2 the information being free from signals that indicate the existence of the input devices, or display;
2.1.3 a wide range of interactions based on movement of the subject's sensory organs with the operator able to see the effect of moving his/her limbs in the VE;
2.1.4 a high correlation between movements of the operator sensed directly and the actions of the representation of the person in the VE;
2.1.5 an ability to change the virtual environment;
2.1.6 a similarity in visual appearance of the subject and their representation in the VE, including the subject's identification between the their own body with that of the representation;
2.1.7 adaptation through learning over time and thus an increase of subject familiarisation with the relationship between motor actions, controls and feedback through the input channels to the senses;
2.1.8 that the feedback loop from operator (use of motor effectors) to system to operator (the effects experienced by the operator as a result of operator actions) form a consistent and lawful whole (that is afference is lawfully related to efference).
2.1.9 that the linkage between afference and efference be simple enough for the subject to be able to model the system over time;
2.1.10 objects in the VE that spontaneously respond to the subject.

The human participant in a VE does not simply soak up everything offered to his senses - the inputs are mediated through the fidelity of the sense organs, the mental models and representation systems typically employed by the particular person involved when interpreting and acting in everyday reality. For example, a high bandwidth and high resolution visual input showing a great deal of visual detail and colour might be very important in increasing the sense of presence for a person who notices and reacts to such detail in everyday reality, but of no importance to someone who usually only notices and typically reacts to just the vaguest outlines of things in her immediate environment, and who pays more attention to sounds and tactile sensations. A VEG relying largely on visual signals (as systems commonly do today) may induce widely differing responses from users who process inputs in quite different ways from each other.

A VEG ideally provides input to the human's visual, auditory, kinesthetic (tactile and haptic) (and maybe eventually olfactory and gustatory) sense organs. The system also provides the possibility of an immersed, first person experience of the virtual environment. In searching for a

suitable psychological approach appropriate for understanding the human's response to the VE, we came across the rather unorthodox therapeutic model known as Neuro-Linguistic Programming (NLP) (see, for example, Dilts, 1979). This seemed relevant because it is based on the two central ideas of *representation systems* and *perceptual position.*

This model claims that subjective experience is encoded in terms of three main representation systems, Visual, Auditory and Kinesthetic (VAK). The Visual system includes external images and remembered and constructed internal images. The Auditory system includes external sounds, and internal remembered and constructed sounds. It also includes internal dialogue, that is the person talking to himself on the inside. The Kinesthetic system includes tactile sensations, the haptic sensations caused by external forces acting on the body, and also emotional responses (which are decomposed into specific patterns of internal tactile and haptic sensations). It is claimed that people have a tendency to prefer one representation system over the others, at least in a given context - for example, a person may reason out problems by thinking mainly in images, whereas another may use mainly internal dialogue (the so-called, auditory-digital system). It is a hypothesis of NLP that the predicates (Visual, Auditory, Kinesthetic) a person uses reflect the underlying pattern of their thought processes - that a V predicate actually indicates visual thinking, for example.

The perceptual position is the standpoint from which the person experiences and remembers events. A person might remember an event from an associated (egocentric) standpoint, and see the event unfolding in his mind's eye from the viewpoint in which he originally experienced it. This is called the first perceptual position. Alternatively a person might remember the event from a dissociated (exocentric) perspective - either from the point of view of another actor in the scene (second position), or from an abstract, disembodied point of view (third position). For example, a person trying to convince someone in an argument might say: "I can feel that it is right" (first position, K) or "You can tell that it is right" (second position, A) or "It can be seen that it is right" (third position, V). The representation systems and perceptual position are logically orthogonal - there being nine possible combinations in this example.

The VEG that we are using presents mainly visual information to the participant, with a small amount of (poor quality) auditory information, and no tactile feedback. We formed the hypothesis that people whose preferred representation system is Visual are likely to have a higher degree of presence (other things being equal of course) than those whose primary system is auditory or kinesthetic. Similarly, people who process information primarily from the first perceptual position, are more likely to experience a higher degree of presence than those from second or third. The K system is more complex, since although there is no K feedback to the participants, nevertheless they must carry out physical movements that are similar to those necessary in real life to achieve a similar outcome - turning their heads to change their direction of gaze, bending down and moving their hands in a realistic manner to pick up objects. Since participants operate in the VE in something like the way that they operate in everyday reality, an environment that offers such a possibility may provide an enhanced sense of presence. On the other hand, they must carry out certain kinesthetic actions in a way quite contrary to everyday experience - such as moving forwards by pressing a button on a 3D mouse, and navigating by hand orientation.

We are not taking a standpoint here on the validity of the NLP claims, but rather using the model as a basis for the formation of testable hypotheses of interest in our VE research. This model seemed to be particularly interesting, since its major variables (representation systems and perceptual position) are related to precisely those factors distinguishing IVEs from other forms of human-computer interaction - the provision of a first person experience based on inputs to the three major sensory channels.

3. The Pilot Experiment

The experiments described in this paper were implemented on a DIVISION ProVision system (Grimsdale, 1991). The VEG itself consisted of a standard ProVision system, a DIVISION 3D mouse (the input device), and a Virtual Research Flight Helmet™ as the head mounted display (HMD). A flock of two Ascension Birds™ is used for position tracking of the head and the

mouse. The images are generated using Intel i860 microprocessors (one per eye) to create an RGB RS-170 video signal which is fed to an internal NTSC video encoder and then to the displays of the Flight Helmet™. These displays (for the left and right eye) are colour LCDs with a 360×240 resolution and the HMD provides a field of view of about 100° along the horizontal with a consequent loss of peripheral vision. During the experiment, the update rate varied between about 6 and 16 frames per second.

For our pilot study on presence twenty students studying human computer interaction on a Masters program were invited to take part as subjects. They were told nothing about the purposes of the study, except that it involved "virtual reality". They were divided into two groups - roughly matched with respect to sex and whether or not they were native English speakers, and half assigned to a control group and the other half to the experimental group. Finally seventeen students actually participated in the study, nine of whom were in the experimental group. The control group subjects were endowed with a disembodied three dimensional arrow cursor, that responded correctly to hand movements with the 3D mouse in terms of orientation. The experimental group had a virtual body with the right hand and arm slaved to movements of the subject's real right hand holding the 3D mouse, and the orientation of the rest of the body slaved to the subject's gross body movements.

All subjects were presented with the identical scenario - a corridor with six doors each leading to a room that exercised a feature of the VE-person interaction that we wished to examine. A full description of the scenarios can be found in (Slater and Usoh, 1992). The subjects were given a questionnaire to complete, Part A immediately after the experience, and Part B part 24 hours later. In Part A they answered questions about their reported sense of presence, physical and mental responses, and basic facts such as their handedness and eyesight. Part B asked them to recall their experiences and write about their sense of presence and anything else they wanted to say about the experience.

Table 1
Reported Sense of Presence

To what extent did you experience a sense of being "really there" inside the virtual environment?	No.
(1) not at all really there	1
(2) there to a small extent	2
(3) there to some extent	5
(4) a definite sense of being there	3
(5) a strong experience of being there	5
(6) totally there	1

Presence was assessed in several ways, but here we consider only the reported degree of presence, with the responses shown in Table 1. The idea of asking the subjects "Where are you?", although obvious now, did not occur to us before these experiments, but was suggested in the essay written by one of them after the experience. She wrote: "Looking back it feels more like somewhere I visited, rather than something I saw (as in a film), so I suppose I must have felt I was in the scene. I did feel quite immersed in it at the time, but was aware also that at the same time I was standing in the corner of a room with a helmet on."

4. Some Results Relating to Exogenous Factors

We asked two questions on the factors that increased or decreased the sense of presence and the responses are summarised in Table 2.

Amongst the factors that increased the sense of presence, the greatest number of people mentioned the last room entered in the experiment, where they were asked to stand on a plank over a precipice, and drop an object off the side. Even some individuals who reported a relatively low sense of presence said that their presence increased when faced with the plank room, leading us to believe that people generally report their average rather than modal sense of presence.

It is interesting to speculate about how the overall results would have changed if the plank room had been the first room entered by the subjects rather than the last. Assuming that this increased sense of danger led to an increase in the sense of presence, we suspect that the overall sense of presence would have been maintained at a higher level throughout. We have some evidence to suggest that this is the case based on subsequent experimentation.

Table 2
Factors Affecting the Sense of Presence
(Classification of Open-Ended Questions)
(The subject could give more than one response)

(a) Were there any circumstances that especially *increased* your sense of being "really there"? If so write them down, or else write "NONE".	No. out of 17	(b) Were there any circumstances that especially *decreased* your sense of being "really there"? If so write them down, or else write "NONE".	No. out of 17
1. being able to move around	6	1. outside events (including instructor)	4
2. interacting with objects/doing a task	7	2.screen/updates/lag/resolution	6
3. great concentration	1	3. things don't behave naturally (laws of physics are violated)	10
4. mention of body (exp. group only)	3	4. things aren't done naturally	9
5. being on plank ("fear reaction")	11	5. body doesn't behave naturally (exp. group only)	3
6. being upside down	1		

The largest group of subjects mentioned, as something that decreased the sense of presence, the fact that the VE did not behave according to the laws of physics, and that activities were not carried out "naturally" - for example, navigating the environment with the aid of a 3D mouse. In our VE to date, objects are not solid - so that people can walk through walls, and place their virtual hand inside objects. It suggests that people are expecting "reality", and when they don't find it, this causes them to note the disjunction between their expectations and what they experience, which at least temporarily decreases the sense of presence. Gravity in our system at the moment only works partially - objects have no centre of gravity, so that when they fall, if they happen to land on a corner or an edge they will remain balanced on that edge. Some subjects noted their disappointment that the object dropped over the precipice did not shatter when hitting the floor below - and also there was no corresponding sound.

The point about consistent and lawful behaviour of the VE relates to 2.1.1, 2.1.8 and 2.1.9 amongst the factors thought likely to contribute to presence. The subjects also spontaneously mentioned factors that lend weight to points 2.1.4 (for some of those with the virtual body), 2.1.5 - some subjects mentioning that the construction of a pile of blocks in one of the rooms increased their sense of presence. We would have to add to the list the following:

- The ability of subjects to be able to move around in the VE (this is perhaps implicit in 2.1.3);
- The presentation of the subjects with a scenario or events that if presented in everyday reality would normally result in an assessment of their personal relationship to those events, in particular the implications for their own state of mind or body. Facing the subject with a potential danger is one way of realising this that we have used to date.

5. Some Results Relating to Endogenous Factors

Although this was not planned in advance, we made use of the essays that were written by the subjects 24 hours after the experiments. They answered the following question: "Write as much as you want about your overall experiences in the virtual environments. Pay attention to your sense of being there or not, your physical sensations, your mental experiences, your thoughts about what happened - in fact about anything that occurs to you about what you experienced." The responses were short essays ranging in length from 4 to 31 sentences.

We classified each sentence according to its perceptual position, and counted the numbers (if any) of Visual, Auditory, and Kinesthetic predicates and references used. For example, one

subject wrote: "Sensation is similar to being in a dream you know is a dream. Like you're there but you know it's not real." These would be classified as "second position" sentences, with no clue as to the representation system ("sensation" is neutral, since it may be a V, A, or K sensation that the subject is referring to). Another subject wrote: "In many of the rooms I visited, I felt I was really in that world". This would be classified as first perceptual position, with a K predicate used ("felt"). As a last example, someone wrote: "The quality of the image was very grainy, so required a lot of concentration in looking around." This would be classified as third position, with a visual reference ("looking around").

The independent variables were the V,A,K, first, second and third position counts as a proportion of the number of sentences. The experimental control variable, the virtual body or arrow pointer was an additional independent variable. The dependent variable was the reported level of presence on the scale of 1 to 6. Since this was an exploratory analysis, we took the liberty of using this ordered categorical variable as if it were measured on an interval scale based on a sound metric. The justification can only be the utility of the obtained results in generating interesting hypotheses for our later research.

We used multiple regression analysis to fit linear models relating the dependent variable to the independent variables. The advantage of multiple regression is that it takes into account the influence of many independent variables at once, unlike cross tabulations where this is impossible without a very large data set.

There is not the space here to go into an explanation of our model fitting strategy nor the details of the results, which are reported in full in (Slater and Usoh, 1993b). The results were quite startling. We found a very high (statistically significant) multiple correlation coefficient, and such that the fit would be significantly worsened with the deletion of any independent variable. The regression model suggested the following hypotheses:

(a) That independently of whether or not the subject has a virtual body, the higher the proportion of visual predicates and references used, the greater the sense of presence, and the higher the proportion of auditory predicates and references the lower the sense of presence.
(b) For those with a virtual body, the higher the proportion of kinesthetic references and predicates the higher the sense of presence. For those without a virtual body, the higher the sense of kinesthetic terms the lower the sense of presence.
(c) The level of presence increases with first perceptual position (P1) up to the mean level of P1, and then decreases. (The model was quadratic in P1). This is the same for each group, except that the rate of change is steeper for those in the control group.

The quadratic term accounting for (c) might simply be an artifact of the "measurement" of presence - that the underlying "scale" of presence (if there is such) might not be an affine map of our ordinal scale. On the other hand, it could indicate for example, that presence increases with the extent to which a person generally experiences the world from "first position", but that a person who is that way to an extreme, who never takes a dissociated standpoint, cannot allow that "suspension of disbelief" necessary in order to achieve a degree of presence. This idea remains to be explored.

6. Conclusions

Presence is a rather difficult concept to define and measure, even with respect to the phenomenon in everyday reality. For example, someone waiting at a bus stop in a busy urban setting is thinking (visualising, hearing the sounds, feeling the feelings) of a recent trip to the countryside. Are they "present" at the bus stop, or in the countryside? The acid test is, perhaps, the question "Where are you?" The person would probably say "at the bus stop" - although the question itself has forced them out of their "day dreaming". On the other hand, someone in a trance induced experience of being in the countryside, would probably answer "in the countryside". From this argument we could say that presence in an IVE involves the commitment of the person's entire neurology to the "suspension of disbelief" that they are ""somewhere else" rather than where their

physical body really is, and, in order to answer the question "where are you?" they are not forced to return to everyday reality, and would give an answer indicating the "somewhere else". In the case of presence in an IVE, the somewhere else is computer generated.

In this paper we have attempted to give an overview of our work so far, in generating hypotheses for factors influencing presence in IVEs in the context of interior walkthroughs. Our pilot studies have indicated that the exogenous factors suggested in the literature to date, discussed in Section 2, are important and necessary conditions. A particularly important requirement, at least in our walkthrough application, is that the VE obey the laws of physics in a consistent manner. We also found that the "startle response" induced by a presentation of the subject to an unusual situation possibly involving personal danger, probably increases the level of presence, and therefore cannot by itself be used as a measurement of presence. We have used the ideas of neuro-linguistic programming as a first attempt to elicit subjective endogenous factors that might help explain people's differing responses to their VE experiences. The pilot results suggest that representation systems and perceptual position do help explain the variation in responses to the system used in our experiments. However, causality cannot be inferred, since the essays from which the data were generated were written after the experiences to which they refer.

Acknowledgements

We would like to thank Anthony Steed for helping with the conduct of the experiments and all those who participated in the experiments, in particular the 1992-3 Human-Computer Interaction stream of the MSc in Computer Science. This work is funded by the London Parallel Applications Centre, in a collaborative project with DIVISION UK Ltd, and Thorn EMI Central Research Laboratories.

References

Brooks, F.P. et. al. (1992) Final Technical Report: Walkthrough Project, Six Generations of Building Walkthroughs, Department of Computer Science, University of North Carolina, Chapel Hill, N.C. 27599-3175.

Dilts, R., J. Grinder, R. Bandler, J. DeLozier, L. Cameran-Bandler (1979) Neuro-Linguistic Programming I, Meta Publications.

Ellis, S.R. (1991) Nature and Origin of Virtual Environments: A Bibliographic Essay, Computing Systems in Engineering, 2(4), 321-347.

Grimsdale, C. (1991) dVS-Distributed Virtual Environment System, Proceedings of Computer Graphics 91 Online Conference, London.

Harriss, P. and C. McKellar (1992) Real Time Lighting for Virtual Worlds, Virtual Reality International '92 Olympia (Meckler).

Heeter, C. (1992) Being There: The Subjective Experience of Presence, Telepresence, Presence: Teleoperators and Virtual Environments, 1(2), spring 1992, MIT Press, 262-271.

Held, R.M. and N.I. Durlach (1992) Telepresence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 109-112.

Loomis, J.M. (1992a) Distal Attribution and Presence, Telepresence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press, 113-119.

Loomis, J.M. (1992b) Presence and Distal Attribution: Phenomenology, determinants, and assessment, SPIE 1666 Human Vision, Visual Processing and Digital Display III, 590-594.

Sheridan, T.B. (1992) Musings on Telepresence and Virtual Presence, Telepresence, Presence: Teleoperators and Virtual Environments, 1, winter 1992, MIT Press,120-126.

Slater, M. and M. Usoh (1992) An Experimental Exploration of Presence in Virtual Environments, Department of Computer Science, QMW University of London, report.

Slater, M. and M. Usoh (1993a) The Influence of a Virtual Body on Presence in Immersive Virtual Environments, VR 93, Virtual Reality International, Proceedings of he third annual conference on Virtual Reality, London, pp34-42. Meckler.

Slater, M. and M. Usoh (1993b) Representation, Perceptual Position and Presence in Virtual Environments, Department of Computer Science, QMW University of London, submitted.